

Detecting Late Payments by Students with Random Forest and Particle Swarm Optimisation

Abdul Hamid¹, Sugiono^{*2}, Henny Destiana³, Titin Prihatin⁴, Astrilyana⁵

^{1,2,3,4,5}Universitas Bina Sarana Informatika, Jakarta, Indonesia

^{*}Corresponding Author: sugiono.sgx@bsi.ac.id

Abstract - Delays in tuition fee payments are a crucial problem commonly faced by private universities in Indonesia. This problem not only complicates campus financial management, but also has the potential to hinder the smooth running of student studies, such as leave of absence or discontinuation of studies. To date, there are not many predictive systems used to detect potential late payments early, especially the Random Forest (RF) model with Particle Swarm Optimisation (PSO). Therefore, this study aims to develop a predictive model for student payment delays by utilising the RF algorithm optimised using the PSO feature selection method. The Dataset used consists of 15,697 student data covering academic and administrative attributes. Pre-processing was carried out to convert categorical data into numerical form so that it could be processed by the classification algorithm. The evaluation results show that the RF model without optimisation produces an accuracy of 97.37%, precision of 100%, recall of 18.68%, and AUC of 0.825 ± 0.020 . After feature selection with PSO, the model performance improved, with an accuracy of 98.83%, precision of 98.20%, recall of 25.40%, and AUC remaining stable at 0.825 ± 0.035 . The most influential attributes in the classification were semester, leave status, studying while working, and father's occupation. The results of this study indicate that the combination of RF and PSO can produce an efficient and accurate prediction model, which can be used as a decision-making tool in higher education administration management.

Keywords— Late Payments, Classification, PSO, Random Forest, Feature Selection.

I. PENDAHULUAN

Pendidikan tinggi swasta di Indonesia menghadapi tantangan dalam memastikan keberlangsungan studi mahasiswa secara akademik dan administratif [1]. Salah satu persoalan krusial adalah keterlambatan pembayaran biaya pendidikan, yang kerap berdampak pada status akademik mahasiswa, termasuk kemungkinan mengambil cuti kuliah bahkan berhenti sementara dari studi [2]. Keterlambatan ini tidak hanya menyulitkan pihak administrasi kampus dalam perencanaan keuangan, tetapi juga berpotensi mengganggu kelancaran akademik mahasiswa yang bersangkutan [3].

Permasalahan tersebut mendorong perlunya sistem prediktif yang dapat mendeteksi keterlambatan pembayaran sejak dini [4]. Dengan mengetahui potensi keterlambatan mahasiswa, kampus dapat melakukan tindakan mitigasi seperti pemberian pengingat, konseling keuangan, atau skema pembayaran alternatif sebelum mahasiswa mengalami hambatan akademik [5].

Berbagai penelitian sebelumnya telah mencoba mengklasifikasikan atau memprediksi keterlambatan pembayaran menggunakan pendekatan data mining. Salah satu penelitian menggunakan algoritma Random Forest (RF) untuk memprediksi kelancaran pembayaran mahasiswa dan menunjukkan performa klasifikasi yang baik [6]. Penelitian lain mengimplementasikan algoritma K-Nearest Neighbor (KNN) dan Naive Bayes untuk klasifikasi keterlambatan pembayaran [7], sementara perbandingan performa antara Naive Bayes dan Support Vector Machine juga telah dilakukan [8]. Pendekatan klasifikasi reguler dengan KNN tanpa optimasi turut digunakan dalam penelitian lain [9]. Selain itu, terdapat studi yang memfokuskan pada prediksi kemungkinan cuti studi mahasiswa dengan algoritma C4.5 [10].

Beberapa penelitian juga mulai melibatkan optimasi fitur untuk meningkatkan akurasi model [11]. Metode seleksi fitur dibandingkan dalam konteks prediksi mahasiswa yang berisiko dropout [12]. Penelitian lainnya menggabungkan algoritma Support Vector Machine dengan Particle Swarm Optimization (PSO) untuk mendeteksi keterlambatan pembayaran, dan menunjukkan peningkatan akurasi dibandingkan model tanpa optimasi [13].

Berdasarkan celah penelitian sebelumnya, studi ini mengusulkan pengembangan model prediksi keterlambatan pembayaran mahasiswa menggunakan algoritma RF yang dioptimasi dengan PSO. RF dipilih karena kemampuannya menangani data yang kompleks dan tidak linear, serta memberikan interpretabilitas terhadap fitur yang memengaruhi klasifikasi [14]. Sementara itu, PSO digunakan untuk memilih subset fitur yang paling relevan sehingga dapat meningkatkan akurasi model dan efisiensi komputasi [15].

DOI: <https://doi.org/10.24036/voteknika.v13i4.135375>

Received : 2025-08-07
Revised : 2025-11-14
Accepted : 2025-11-19
Published : 2025-12-29



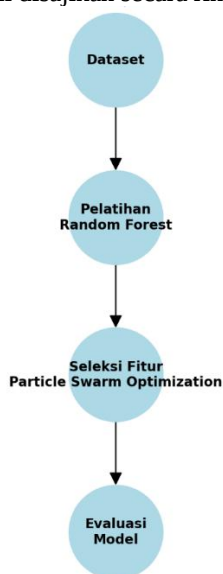
For all articles published in VOTETEKNIKA
<https://ejournal.unp.ac.id/index.php/voteknika>, © copyright is retained by the authors. This is an open-access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Penelitian ini bertujuan untuk menghasilkan model prediksi yang akurat dan adaptif dalam mendeteksi mahasiswa yang berpotensi terlambat membayar, guna mendukung pengambilan keputusan administratif oleh pihak kampus [16]. Kontribusi utama dari penelitian ini adalah penerapan kombinasi RF dan PSO dalam konteks prediksi keterlambatan pembayaran, yang masih jarang dieksplorasi. Berbeda dengan penelitian yang hanya menggunakan RF tanpa optimasi [6], serta yang hanya menggabungkan SVM dengan PSO [13], penelitian ini menghadirkan pendekatan baru yang mengoptimalkan RF dengan PSO, sehingga menjadi lebih efisien dan akurat [17].

Dengan demikian, penelitian ini tidak hanya berkontribusi dalam pengembangan sistem deteksi keterlambatan pembayaran, tetapi juga memberikan kebaruan metodologis dalam optimalisasi algoritma klasifikasi berbasis ensemble untuk konteks manajemen pendidikan tinggi..

II. METODE

Penelitian ini disusun melalui serangkaian tahapan terstruktur untuk membentuk model prediksi keterlambatan pembayaran mahasiswa, yang bertujuan membantu pencegahan potensi cuti atau putus studi. Proses penelitian mencakup beberapa langkah inti, dimulai dari pengumpulan data, dilanjutkan dengan pelatihan RF, kemudian seleksi fitur menggunakan PSO. Tahap inti tersebut mempunyai rincian dengan pengumpulan data dari sistem informasi akademik, kemudian dilanjutkan dengan tahap pra-pemrosesan, yang mencakup pengkodean data kategorikal, normalisasi, serta pembersihan data dari nilai hilang maupun inkonsistensi. Setelah itu, data dibagi ke dalam subset pelatihan dan pengujian melalui *k-fold cross-validation*. Pada tahap selanjutnya, dilakukan pelatihan model RF sebagai *baseline*. Setelah *baseline* terbentuk, dilakukan seleksi fitur menggunakan PSO yang berperan sebagai metode wrapper dengan RF sebagai evaluator. Fitur terbaik hasil seleksi PSO kemudian digunakan untuk melatih ulang model RF sehingga performanya dapat dibandingkan dengan *baseline*. Tahap akhir penelitian meliputi evaluasi model menggunakan metrik seperti *confusion matrix*, akurasi, dan AUC. Alur lengkap tahapan penelitian ini disajikan secara ringkas pada Gambar 1.



Gambar. 1 Langkah Penelitian

A. Dataset

Penelitian ini diawali dengan pemanfaatan *Dataset* sebagai fondasi utama dalam proses pembentukan model prediksi. *Dataset* yang digunakan diperoleh dari salah satu perguruan tinggi swasta di Indonesia. Data ini terdiri dari 15.697 *record* mahasiswa yang diambil dari sistem informasi akademik. Setiap *record* merepresentasikan informasi administratif dan akademik mahasiswa yang relevan untuk proses klasifikasi status pembayaran.

Tabel 1. Deskripsi data.

Rule	Nama	Tipe	Keterangan	Nilai/encoding
label	no_by_klh	binomial	nomor biaya kuliah untuk mengidentifikasi mahasiswa telat bayar atau tidak	0 telat bayar, 1 untuk lunas
reguler	smt	integer	semester aktif	1 sampai 8
reguler	jen_kel	integer	jenis kelamin	0 pria, 1 wanita
reguler	Jrs_slta	integer	jurusan sesuai atau tidak pada saat di SLTA atau sederajat	0 untuk jurusan tidak sesuai dan 1 untuk sesuai
reguler	klh_smbal_krj	integer	mahasiswa sambal bekerja atau tidak	0 untuk kuliah tidak bekerja, 1 untuk kuliah sambal kerja
reguler	pkj_ayah	integer	jenis pekerjaan ayah	Berbagai jenis pekerjaan ayah
reguler	sts_aktif_cuti	integer	pernah cuti	0 untuk tidak pernah cuti, 1 pernah cuti

Tabel 1 Terdapat 7 atribut dalam *Dataset* ini, dengan satu di antaranya berfungsi sebagai atribut target (label), yaitu status pembayaran mahasiswa (tepat waktu atau terlambat). Enam atribut lainnya merupakan variabel input yang digunakan sebagai prediktor.

Pada penelitian ini, proses pra-pemrosesan dilakukan menggunakan perangkat lunak *RapidMiner*, yang menyediakan berbagai operator otomatis untuk transformasi data. *Dataset* yang digunakan tidak mengandung nilai kosong, sehingga tidak diperlukan proses imputasi atau penggantian nilai hilang. Namun, terdapat beberapa penyesuaian terhadap format dan representasi nilai untuk meningkatkan kejelasan dan keterbacaan data.

B. Pelatihan Model RF

RF termasuk ensemble learning yang digunakan untuk klasifikasi. Algoritma ini bekerja dengan cara membangun sejumlah pohon keputusan secara acak pada saat pelatihan dan menggabungkan hasilnya untuk mendapatkan prediksi akhir yang lebih akurat dan stabil dibandingkan dengan menggunakan satu pohon keputusan saja. Dalam penelitian ini, pelatihan model RF dilakukan dengan menggunakan data pelatihan yang telah melalui proses pra-pemrosesan, termasuk normalisasi dan encoding variabel kategori. Parameter utama yang diatur pada model ini meliputi jumlah pohon keputusan,

kedalaman maksimum pohon, serta jumlah fitur yang dipertimbangkan pada setiap split node.

C. Seleksi Fitur dengan PSO

Seleksi fitur merupakan tahap penting dalam pengolahan data untuk meningkatkan performa model serta mengurangi kompleksitas komputasi. Dengan memilih fitur yang paling relevan, model dapat lebih fokus pada variabel yang berkontribusi signifikan terhadap prediksi. Pemilihan fitur yang digunakan untuk itu yaitu PSO. PSO adalah algoritma optimasi berbasis populasi dengan menggerakkan sejumlah partikel pada saat pencarian, sehingga menemukan fitur yang berbobot untuk memaksimalkan model RF.

D. Evaluasi Model

Evaluasi model dilakukan dalam dua tahap, yaitu sebelum dan sesudah diterapkannya seleksi fitur menggunakan *Particle Swarm Optimization* (PSO) pada algoritma Random Forest. Tujuannya adalah untuk mengetahui sejauh mana pengaruh seleksi fitur terhadap performa model dalam melakukan klasifikasi.

Evaluasi dilakukan menggunakan data uji yang sama pada kedua model berdasarkan hasil dari *confusion matrix* dan AUC yang didapat dari kedua metode tersebut. Hasil dari *confusion matrix* dihitung dengan rumus berikut:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Presisi = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Pada rumus tersebut dimana TP merupakan kepanjangan dari *True Positive*, TN merupakan kepanjangan dari *True Negative*, FP merupakan kepanjangan dari *False Positive*, FN merupakan kepanjangan dari *False Negative*.

III. HASIL DAN PEMBAHASAN

A. Dataset

Dataset yang digunakan dalam penelitian ini berasal dari sistem informasi akademik salah satu perguruan tinggi swasta di Indonesia, dengan total sebanyak 15.697 record mahasiswa. Data mencakup berbagai atribut yang berkaitan dengan kondisi akademik dan non-akademik mahasiswa, yang berpotensi memengaruhi kelancaran pembayaran.

Pada tahap awal, data masih berada dalam bentuk mentah dengan format kategorikal teks. Data sebelum dilakukan pra-pemrosesan dengan beberapa atribut masih menggunakan deskripsi verbal, seperti "Laki-Laki" untuk jenis kelamin, "Sesuai jurusan" untuk jurusan SLTA, "Ya" atau "Tidak" untuk kuliah sambil bekerja, serta "Petani" atau "Wiraswasta" atau "Lainnya" untuk pekerjaan orang tua. Untuk keperluan klasifikasi dengan algoritma Random Forest, seluruh atribut teks tersebut dikonversi ke bentuk numerik. Proses konversi dilakukan dengan Laki-laki menjadi 0, perempuan menjadi 1, sesuai jurusan di ubah menjadi 1, tidak sesuai jurusan di ubah menjadi 0, bekerja diubah menjadi 1, tidak bekerja diubah menjadi 0, pekerjaan ayah dikategorikan menjadi numerik. Hasil akhir dari proses pra-pemrosesan ditampilkan pada Gambar 3, yang menunjukkan data sudah dalam bentuk numerik sepenuhnya dan siap digunakan untuk proses klasifikasi.

ExampleSet (15697 examples, 1 special attribute, 6 regular attributes)

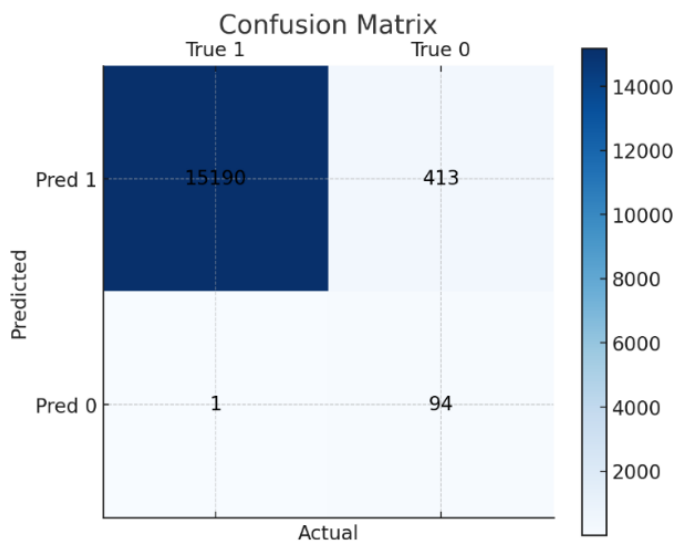
Row No.	no_by_klh	smt	jen_kel	jrs_slta	klh_smbil_krj	pkj_ayh	sts_aktif_cuti
1	1	8	1	1	1	18	0
2	1	8	1	1	1	18	0
3	1	6	0	0	1	18	0
4	1	6	1	0	1	18	0
5	1	6	0	0	1	18	0
6	1	6	0	0	1	18	0
7	1	6	1	0	1	18	0
8	0	4	1	0	1	18	1
9	0	4	0	0	1	18	1
10	1	6	0	0	1	18	0
11	1	6	1	0	1	18	0
12	1	6	0	0	1	18	0
13	1	6	0	0	1	18	0
14	1	6	1	0	1	18	0
15	1	6	1	0	1	18	0
16	1	4	1	0	1	18	0
17	1	2	1	0	1	16	0
18	1	2	0	0	1	16	0
19	1	2	0	0	0	16	0
20	1	2	0	0	0	16	0
21	1	2	0	0	1	18	0
22	1	2	0	0	0	16	0
23	1	2	0	0	0	16	0

Gambar 2. *Dataset* Pembayaran

Gambar 2 memperlihatkan *sample Dataset* setelah melalui tahap pra-pemrosesan. Setiap baris merepresentasikan satu mahasiswa, sedangkan setiap kolom menunjukkan nilai dari atribut-atribut yang telah disesuaikan dalam bentuk numerik. Data yang telah dibersihkan dan ditransformasikan ini selanjutnya digunakan dalam proses pelatihan dan pengujian model klasifikasi. Dengan struktur yang sudah seragam secara numerik, algoritma RF dapat membaca pola-pola dalam data secara lebih optimal dan akurat, terutama dalam mengidentifikasi mahasiswa yang berpotensi mengalami keterlambatan pembayaran yang berdampak pada status akademik.

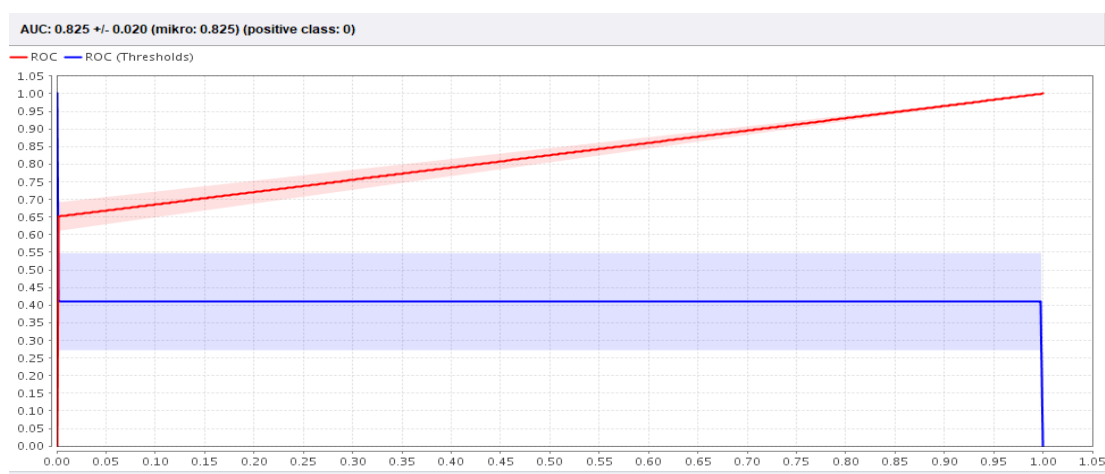
B. Pelatihan Model RF

Model RF dilakukan dengan memproses data pembayaran yang telah selesai dari pra-pemrosesan data. Hasil evaluasi model klasifikasi menggunakan RF disajikan dalam bentuk *confusion matrix* Gambar 3.



. Gambar 3. Confusion Matrix RF

Berdasarkan confusion matrix Gambar 3, diperoleh nilai-nilai berikut: Nilai TP 15.190 mengindikasikan bahwa mahasiswa yang mengalami keterlambatan benar terklasifikasi sebagai mahasiswa yang mengalami keterlambatan pembayaran. FP 413 mengindikasikan bahwa mahasiswa yang tidak terlambat namun salah diklasifikasi sebagai terlambat dalam pembayaran. TN 94 mahasiswa yang tidak terlambat, dan benar diklasifikasi tidak terlambat. FN 1 mahasiswa yang terlambat, tapi diklasifikasi tidak terlambat. Hasil dari *confusion matrix* tersebut setelah dihitung akan menghasilkan nilai akurasi sebesar 97,37%, precision 100,00%, 18,68%. Sebagai pelengkap pelatihan model RF selain *confusion matrix* menghasilkan juga kurvo roc yang berfungsi untuk kemampuan model dalam membedakan antara mahasiswa yang terlambat pembayaran dan mahasiswa yang tidak terlambat pembayaran. Visualisasi kurva ROC dari model RF disajikan pada Gambar 4 berikut.



Gambar. 4 Kurva ROC Random Forest

Pada Gambar 4, menunjukkan performa model yang cukup baik dalam membedakan antara mahasiswa yang terlambat dan tidak terlambat dalam melakukan pembayaran. Hal ini ditunjukkan oleh nilai Area Under Curve (AUC) sebesar 0,825, yang mengindikasikan bahwa model memiliki kemampuan klasifikasi yang tergolong kuat. Dengan nilai AUC di atas 0,8, dan penyimpangan standar (standard deviation) 0,035 model RF dapat dikategorikan memiliki performa klasifikasi yang andal dalam konteks prediksi keterlambatan pembayaran mahasiswa. Namun dari nilai-nilai tersebut dapat ditingkatkan dengan metode seleksi fitur dengan PSO.

C. Seleksi Fitur dengan PSO

Meskipun model RF menunjukkan performa yang baik dalam mendeteksi keterlambatan pembayaran mahasiswa, nilai akurasi dan AUC yang diperoleh masih memiliki ruang untuk peningkatan. Salah satu pendekatan yang dapat digunakan untuk meningkatkan kinerja model adalah melalui proses seleksi fitur. Dengan menerapkan PSO, model RF dapat berfokus pada fitur-fitur yang paling informatif. Setelah dilakukan proses seleksi fitur menggunakan PSO, diperoleh bobot untuk masing-masing atribut pada Tabel 2.

Tabel 2. Bobot atribut model PSO

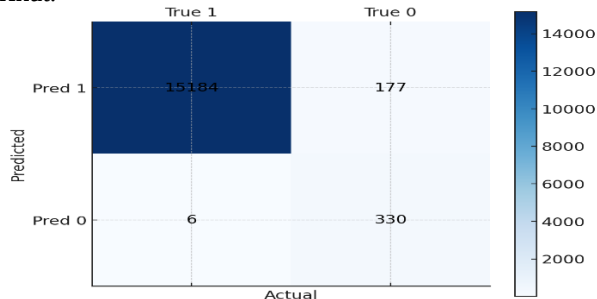
atribut	bobot
semester	1
jenis kelamin	0
sesuai jurusan slta	0
kuliah sambil bekerja	0.263
pekerjaan ayah	0.202
status cuti	0.440

Nilai bobot pada tabel 2 menunjukkan tingkat kepentingan atau kontribusi masing-masing atribut terhadap klasifikasi keterlambatan pembayaran. Tinggi bobot dengan nilai 1 semakin besar pengaruh atribut tersebut dalam menentukan apakah seorang mahasiswa akan terlambat dalam melakukan pembayaran atau tidak. Tinggi bobot dengan nilai 0 maka semakin tidak berpengaruh atribut tersebut dalam menentukan apakah seorang mahasiswa akan terlambat dalam melakukan pembayaran atau tidak.

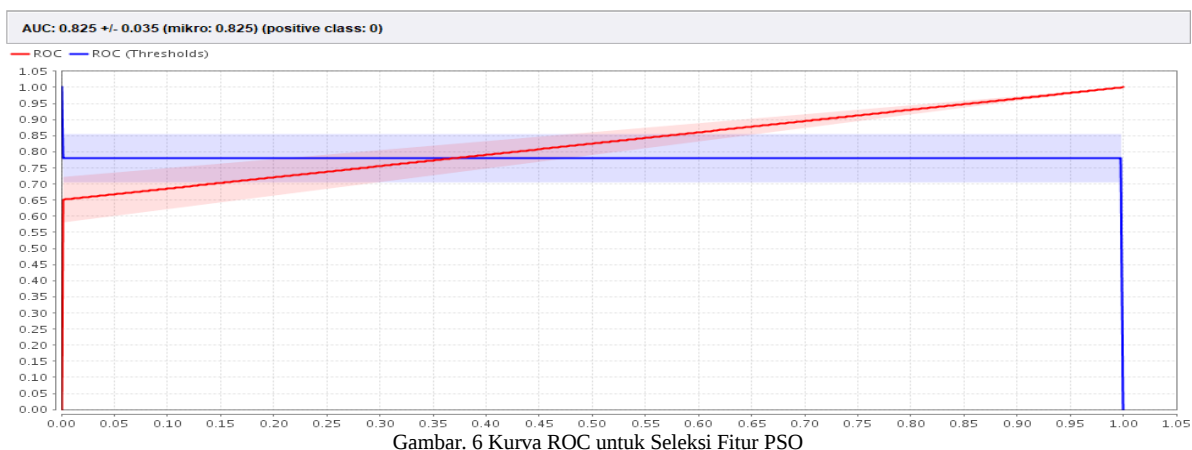
Setelah proses seleksi fitur, hanya atribut-atribut yang memiliki bobot signifikan yang digunakan dalam pelatihan ulang model RF, yaitu semester, kuliah sambil bekerja, pekerjaan ayah, dan status cuti. Sementara itu, atribut seperti jenis kelamin dan kesesuaian jurusan SLTA diabaikan karena

memiliki bobot nol, yang berarti tidak memberikan kontribusi terhadap proses klasifikasi.

Setelah dilakukan pelatihan ulang menggunakan fitur-fitur terpilih tersebut, model kemudian dievaluasi menggunakan data uji yang sama seperti sebelumnya. Hasil evaluasi disajikan dalam bentuk confusion matrix untuk mengetahui peningkatan performa model setelah dilakukan seleksi fitur. *Confusion matrix* setelah optimasi PSO ditampilkan pada Gambar 5 berikut.



Gambar. 5 *Confusion Matrix* hasil Seleksi Fitur PSO



Gambar. 6 Kurva ROC untuk Seleksi Fitur PSO

Kurva ROC pada Gambar 6 menunjukkan peningkatan performa model setelah dilakukan seleksi fitur. Nilai AUC yang dihasilkan adalah sebesar $0,825 \pm 0,035$. Meskipun nilai rata-rata AUC tetap sama dengan model sebelum optimasi (0,825), akan tetapi penyimpangan standar (*standard deviation*) yang lebih besar 0,035 mencerminkan adanya sedikit peningkatan variasi performa antar eksperimen atau fold pada validasi silang.

D. Evaluasi Model

Evaluasi model dilakukan untuk mengukur performa algoritma RF sebelum dan sesudah dilakukan proses seleksi fitur menggunakan PSO, dengan hasil evaluasi Tabel 2 berikut :

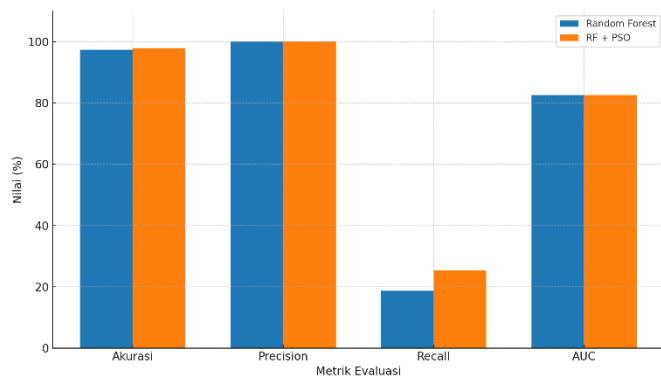
Tabel 3. Hasil Evaluasi Model Random Tree dengan dan tanpa FS

Algoritma	Random Forest	Random Forest + Particel Swarm Optimization
Accuracy	97,37%	98,83%
Precision	100,00%	98,20%
Recall	18,68%	65,09%
AUC	82,50%	82,50%

Gambar 5 matriks ini diperoleh nilai-nilai berikut: Nilai TP 15.184 mengindikasikan bahwa mahasiswa yang mengalami keterlambatan benar terklasifikasi sebagai mahasiswa yang mengalami keterlambatan pembayaran. FP 177 mengindikasikan bahwa mahasiswa yang tidak terlambat namun salah diklasifikasi sebagai terlambat dalam pembayaran. TN 330 mengindikasikan bahwa mahasiswa yang benar terlambat pembayaran diklasifikasi tidak mengalami keterlambatan. FN 6 mengindikasikan bahwa mahasiswa yang mengalami keterlambatan namun salah diklasifikasi tidak terlambat dalam pembayaran. Hasil dari *confusion matrix* tersebut setelah dihitung akan menghasilkan nilai akurasi sebesar 98,83%, precision 98,20%, dan recall 65,09%. Untuk menilai performa model secara lebih mendalam, digunakan kurva ROC Visualisasi kurva ROC dari model PSO disajikan pada Gambar 6.

Dari Tabel 3, Hasil evaluasi model RF setelah seleksi fitur dengan PSO menunjukkan bahwa nilai akurasi, precision, dan recall mengalami peningkatan, disertai penyederhanaan jumlah fitur yang digunakan. Nilai AUC yang diperoleh tetap sebesar 0,825, namun dengan peningkatan standard deviation menjadi $\pm 0,035$. Hal ini menunjukkan bahwa meskipun rata-rata performa tetap, model menjadi sedikit lebih variatif antar fold, namun dengan struktur fitur yang lebih ringkas dan efisien..

Untuk memberikan gambaran yang lebih jelas mengenai perbandingan performa model sebelum dan sesudah seleksi fitur, dibuat visualisasi dalam bentuk diagram batang. Diagram ini membandingkan nilai akurasi, *precision*, *recall*, dan AUC dari model RF dan model RF yang telah dioptimasi dengan PSO terdapat pada Gambar 8.



Gambar. 8 Perbandingan Evaluasi Model

Gambar 8 Secara keseluruhan, hasil evaluasi menunjukkan bahwa kombinasi RF dan PSO mampu memberikan hasil klasifikasi yang kompetitif, dengan efisiensi model yang lebih baik. Pengurangan jumlah fitur tidak menyebabkan penurunan akurasi, melainkan justru mendukung model yang lebih sederhana dan mudah diinterpretasikan, yang sangat berguna dalam konteks pengambilan keputusan administratif di lingkungan perguruan tinggi.

IV. KESIMPULAN

Penelitian ini berhasil membangun model prediksi keterlambatan pembayaran mahasiswa menggunakan algoritma RF yang dioptimasi dengan metode seleksi fitur PSO. Berdasarkan hasil evaluasi, model RF tanpa seleksi fitur menghasilkan nilai akurasi sebesar 97,37%, precision 100%, recall 18,68%. Setelah dilakukan seleksi fitur dengan PSO, model menunjukkan peningkatan pada *recall* menjadi 65,09%, akurasi menjadi 98,83%, dan nilai AUC keduanya 0,825, dengan *standard deviation* masing-masing 0,020 dan 0,035, meskipun dengan penyimpangan standar (*standard deviation*) yang sedikit lebih besar.

Seleksi fitur menggunakan PSO terbukti mampu menyederhanakan model dengan menghilangkan atribut yang tidak relevan, tanpa mengurangi performa secara signifikan. Model ini dapat menjadi alat bantu bagi pihak kampus dalam mendeteksi potensi keterlambatan pembayaran secara dini, sehingga strategi mitigasi seperti pemberian pengingat, konseling keuangan, atau penyesuaian kebijakan pembayaran dapat dilakukan lebih tepat sasaran.

DAFTAR PUSTAKA

- [1] Ridwansyah, M. Iqbal, H. Destiana, Sugiono, and A. Hamid, "Data Mining Berbasis Machine Learning Untuk Analitik Prediktif Dalam Kelulusan," *semantik*, vol. 10, no. 2, pp. 1–10, 2024, doi: <https://doi.org/10.55679/semantik.v10i2.67>.
- [2] S. Anam, M. R. A. Putra, Z. Fitriah, I. Yanti, N. Hidayat, and D. M. Mahanani, "Health Claim Insurance Prediction Using Support Vector Machine With Particle Swarm Optimization," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 17, no. 2, pp. 0797–0806, 2023, doi: [10.30598/barekengvol17iss2pp0797-0806](https://doi.org/10.30598/barekengvol17iss2pp0797-0806).
- [3] N. W. D. Ayuni, N. N. Lasmini, and K. C. Dewi, "Predicting financial distress of property and real estate companies using optimized support vector machine-particle swarm optimization (SVM-PSO)," *Bull. Soc.*

- Informatics Theory Appl.*, vol. 8, no. 1, pp. 97–106, 2024.
- [4] W. Li, "Design of Financial Crisis Early Warning Model Based on PSO-SVM Algorithm," *Math. Probl. Eng.*, pp. 1–8, 2022, doi: <https://doi.org/10.1155/2022/3241802>.
- [5] A. H. Kahfi, T. Prihatin, Yudhistira, A. Sudradjat, and G. Wijaya, "THE RIGHT STEPS TOWARDS GRADUATION: NB-PSO SMART COMBINATION FOR STUDENT GRADUATION PREDICTION," *J. Tek. Inform.*, vol. 5, no. 2, pp. 607–614, 2024, doi: <https://doi.org/10.52436/1.jutif.2024.5.2.1889>.
- [6] T. Azhima, Y. Siswa, and W. J. Pranoto, "Implementasi Seleksi Fitur Information Gain Ratio Pada Algoritma Random Forest Untuk Model Data Klasifikasi Pembayaran Kuliah," *Din. Inform.*, vol. 15, no. 1, pp. 41–49, 2023.
- [7] M. Norhalimi and T. A. Y. Siswa, "Optimasi Seleksi Fitur Information Gain pada Algoritma Naïve Bayes dan K-Nearest Neighbor," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 7, no. 3, pp. 237–255, 2022, doi: [10.14421/jiska.2022.7.3.237-255](https://doi.org/10.14421/jiska.2022.7.3.237-255).
- [8] H. Umar, R. Kusumawati, M. Imamudin, and M. A. Rohman, "Klasifikasi Keterlambatan Pembayaran Sumbangan Pembinaan Pendidikan Menggunakan Algoritma Naïve Bayes dan Support Vector Machine," *TIN Terap. Inform. Nusantara*, vol. 4, no. 11, pp. 709–718, 2024, doi: [10.47065/tin.v4i11.4969](https://doi.org/10.47065/tin.v4i11.4969).
- [9] M. R. Akhmad and T. A. Y. Siswa, "Implementasi K-Nearest Neighbor Dalam Memprediksi Keterlambatan Pembayaran Biaya Kuliah Di Perguruan Tinggi," *Progresif J. Ilm. Komput.*, vol. 18, no. 2, p. 185, 2022, doi: [10.35889/progresif.v18i2.921](https://doi.org/10.35889/progresif.v18i2.921).
- [10] M. Mubarakah, D. Aditya Nugraha, and A. Yunus, "Penerapan Algoritma C4.5 dalam Memprediksi Keterlambatan Pembayaran Sumbangan Pembinaan Pendidikan," *J. Ris. Mhs. Bid. Teknol. Inf.*, vol. 7, no. 1, pp. 1–5, 2024, [Online]. Available: <https://ejournal.unikama.ac.id/index.php/JFTI>.
- [11] T. A. Y. Siswa and R. P. Wibowo, "Komparasi Metode Seleksi Fitur Dalam Prediksi Keterlambatan Pembayaran Biaya Kuliah," *Teknika*, vol. 12, no. 1, pp. 73–82, 2023, doi: [10.34148/teknika.v12i1.601](https://doi.org/10.34148/teknika.v12i1.601).
- [12] M. J. Budiman and Fanny Jouke Doringin, "PENERAPAN ALGORITMA C5.0 DALAM MEMPREDIKSI KETERLAMBATAN PEMBAYARAN BIAYA KULIAH DI UNKRISWINA SUMBA," *J. Ilmu Komput. Revolusioner*, vol. 8, no. 6, 2024.
- [13] H. Nurdin, I. Nawawi, A. Wuryanto, D. Yuliandari, and H. Sugiarto, "Prediksi Keterlambatan Pembayaran Mahasiswa untuk Mitigasi Risiko Cuti Menggunakan SVM Optimasi PSO," *J. Apl. Sains, Inf., Elektron. dan Komput.*, vol. 7, no. 1, pp. 1–11, 2025, doi: <https://doi.org/10.26905/jasiek.v7i1.15483>.
- [14] W. Widiarina, K. Mariskhana, I. D. Sintawati, R. Ningsih, and A. Natong, "Integrating Random Forest and Particle Swarm Optimization for Chronic Kidney Disease Diagnosis," *J. Vocat. Tek. Elektron. dan Inform.*, vol. 13, no. 2, pp. 110–117, 2025, doi: [10.24036/voteteknika.v13i2.133634](https://doi.org/10.24036/voteteknika.v13i2.133634).
- [15] A. Hamid and Ridwansyah, "Optimizing Heart Failure Detection: A Comparison between Naive Bayes and

- Particle Swarm Optimization,” *Paradigma*, vol. 26, no. 1, pp. 30–36, 2024, doi: <https://doi.org/10.31294/p.v26i1.3284>.
- [16] Sumarna, I. Nawawi, Suhardjono, Hari Sugiarto, and D. Yuliandari, “MENINGKATKAN AKURASI PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN METODE ALGORITMA GENETIKA,” *J. Inform. Manaj. dan Komput.*, vol. 16, no. 2, 2024, doi: <http://dx.doi.org/10.36723/juri.v16i2.706>.
- [17] Sumarna, Sartini, W. E. Pangesti, R. Suryadithia, and V. Riyanto, “Decision Tree Optimization in Heart Failure Diagnostics: a Particle Swarm Optimization Approach,” *J. Tek. Inform.*, vol. 5, no. 3, pp. 739–746, 2024, doi: <https://doi.org/10.52436/1.jutif.2024.5.3.1815>.